Roll No. ..... Total Pages : 2

**47420**

**BT-7/M-24**

# DATA SCIENCE WITH R PROGRAMMING

Paper–PC-CS-AIML-401A

Time : Three Hours] [Maximum Marks : 75

**Note :** Attempt *five* questions by selecting at least *one* question from each Unit.

## UNIT–I

**1.** (a) Discuss the benefits of data sciences and big data using suitable examples. (8)

(b) Explain the historical evolution of data science. (7)

**2.** (a) What do you mean by big data ? How it is different from traditional data. (7)

(b) Explain the following data processing operations :
Data cleaning.
Data integration.
Data transformation.
Data reduction. (8)

## UNIT–II

**3.** (a) What do you mean by Data Visualization? Discuss its various types. (8)

(b) Discuss various tools for Data Visualization. (7)

4. (a) Discuss the role of statistics & probability in data science. (7)

(b) Explain the Bayes theorem, how it is useful in calculating conditional probability. (8)

## UNIT–III

5. (a) Discuss basic data types in R programming. (8)

(b) Write a R program to get all prime numbers up to a given number. (7)

6. (a) Write a R program to count number of NA values in a data frame column. (8)

(b) What do you mean by reserved words? Discuss commonly used reserved words in R programming. (7)

## UNIT–IV

7. (a) What do you mean by Linear Regression. Explain its implementation in R programming. (8)

(b) Discuss Hidden Markov Model & its applications in detail. (7)

8. (a) What is the importance of CSV files in machine learning. (7)

(b) Explain the procedure to connect R program to any external interface. (8)

---